DOI: http://dx.doi.org/10.35520/diadorim.2022.v24n1a48566
Recebido em: 05 de dezembro de 2021 /Aceito em: 13 de junho de 2022

SUBJECT EXPRESSION IN SPOKEN VARIETIES OF PORTUGUESE

A EXPRESSÃO DO SUJEITO EM VARIEDADES FALADAS DO PORTUGUÊS

Aline Maria Bazenga[1]

ABSTRACT

Research on spoken varieties of Portuguese allows for the observation of diverse socio-historical and cultural profiles in communities where the language is spoken, and – in the field of linguistic systems – of the occurrence of variable morphosyntactic phenomena. In this article, particular attention is given to the specific phenomenon of variable subject expression. The study is based on a theoretical framework that brings together contributions from discrete theoretical models of Linguistic Variation (Generative Grammar vs. Linguistic Typology). Taken together, the contributions of these models allow for a characterisation of the complexity of the variable expression of the subject in varieties of Portuguese. A third perspective is also added to these two: one in which linguistic variation is related not only to linguistic factors but also to social variation. Within the framework of the model proposed by variationist sociolinguistics, the article seeks to draw a comparison between varieties, to trace the possibilities for a generalisation of null subject and pronominal subject production, to argue for the visualisation of such production along a *continuum*, and to contribute to broadening our understanding of the linguistic phenomenon in question. The results of this comparison of fifteen published works, which were selected due to their common deployment of variationist / quantitative analysis based on *corpora* of Portuguese varieties, points to: i) a highly complex situation in the sphere of Portuguese when non-standard varieties are taken into account, similar to what has been observed in other romance languages such as Italian; ii) the need for a common methodological framework in order to produce a greater quantity of comparable linguistic data; and iii) the need for further theoretical reflection, given the scant attention paid to social or extralinguistic factors.

KEYWORDS: Variable subject expression; Spoken varieties of Portuguese; Variationist and sociolinguistic approaches; Linguistic variation; Null-Subject Languages.

1 Associate Professor in the Department of Languages, Literatures, and Cultures, Faculty of Arts and Humanities, University of Madeira. E-mail: alinebazenga@staff.uma.pt.

A revista ***Diadorim*** utiliza uma Licença Creative Commons - Atribuição-NãoComercial 4.0 Internacional (CC-BY-NC).

RESUMO

A investigação sobre as variedades faladas do português permite-nos observar uma diversidade de perfis histórico-sociais, culturais das suas comunidades, como também, no plano dos sistemas linguísticos, a ocorrência de fenómenos morfossintáticos variáveis. Neste artigo, uma atenção particular será dada a um destes fenómenos: a expressão variável do sujeito. Parte-se de um enquadramento teórico que reúne contributos oriundos de modelos teóricos sobre a Variação Linguística distintos (Gramática Generativa vs Tipologia Linguística), mas cujos contributos permitem configurar o objeto de estudo – a expressão variável do sujeito em variedades do português – sua complexidade. A estes duas perspetivas associa-se uma terceira, na qual a variação linguística está correlacionada não apenas com fatores linguísticos, mas também com a variação social. Pretende-se, no quadro do modelo proposto pela sociolinguística variacionista, estabelecer uma comparação entre variedades, traçar possibilidades de generalização de produção de sujeitos nulos e de sujeitos pronominais, propondo o seu alinhamento em *continua,* e contribuir para o conhecimento mais alargado do fenómeno linguístico em questão. Os resultados, com base em cerca de quinze trabalhos selecionados por terem em comum o facto de recorrerem a análises variacionistas/quantitativas, a partir de *corpora* oral de variedades do português, apontam para (i) uma situação de significativa complexidade na esfera do português, à imagem do observado noutras línguas românicas, como o italiano, quando se têm em conta variedades não padrão, (ii) para a necessidade de um protocolo metodológico em comum, por forma a produzir maior quantidade de dados linguísticos comparáveis e (iii) a necessidade de reflexão teórica, face à pouca relevância observada de fatores sociais ou extralinguísticos.

PALAVRAS-CHAVE: Expressão variável do sujeito; Variedades faladas do português; Abordagens variacionistas e sociolinguísticas; Variação linguística; Línguas de Sujeito Nulo.

Introduction

The diversity within varieties of Portuguese has been widely studied at a lexical and phonetic level and also, although more recently, in the field of syntax, not only from a formal linguistic perspective but also from the starting point of non-formal, functionalist, parametric, and variationist approaches based on empirical data.

The present study seeks to draw a large-scale comparison between the different forms of expressing the subject in varieties of Portuguese, based on the results of studies that have taken non-formal approaches. The Portuguese language data compiled here include not only social and geographical varieties of L1 Portuguese, but also L2 varieties. From the outset, this comparison is a highly complex task (Miestamo, Sinnemäki & Karlsson, 2008; Sampson, Gil & Trudgill, 2009). The variable in question is of a complex nature, as has been frequently observed in the literature on formal and generative linguistics; also, the variety of methodologies adopted in the treatment of empirical data in non-formal studies also complicates this exercise of analysis and comparison of research findings. In short, and as Arends (2001: 180) observes, "more is more complex" – and that is certainly the case in this study.

This article aims to contribute to a better understanding of the variable phenomenon in question across a set of Portuguese varieties for which research data has already been published. The research is divided into two sections, followed by a conclusion. The first section commences with a presentation of the *subject expression* variable within the broader context of interlinguistic variations and takes account of the patterns observed in the Linguistic Typology,

and the most recent discussion of this as a *null subject parameter* within the scope of Generative Grammar. This first section also discusses the variability of subject expression across the group of romance languages within which Portuguese is situated. The second part of the article makes a comparative examination of how the linguistic variable in question functions in different varieties of Portuguese. The analysis proposed here seeks to foreground the convergences and divergences between varieties, insofar as this exercise is possible.

Subject expression in natural languages: a variable linguistic phenomenon

It is widely known that in a language such as Spanish, for example, the subject of a sentence may be expressed overtly as a noun (1a.) or pronoun (1b.); or it may be omitted and not overt, or null (1c.). English, by contrast, allows only the overt expression of the subject, as exemplified in (2).

(1) a. *El niño* estaba estudiando

b. *Yo* estoy cantando

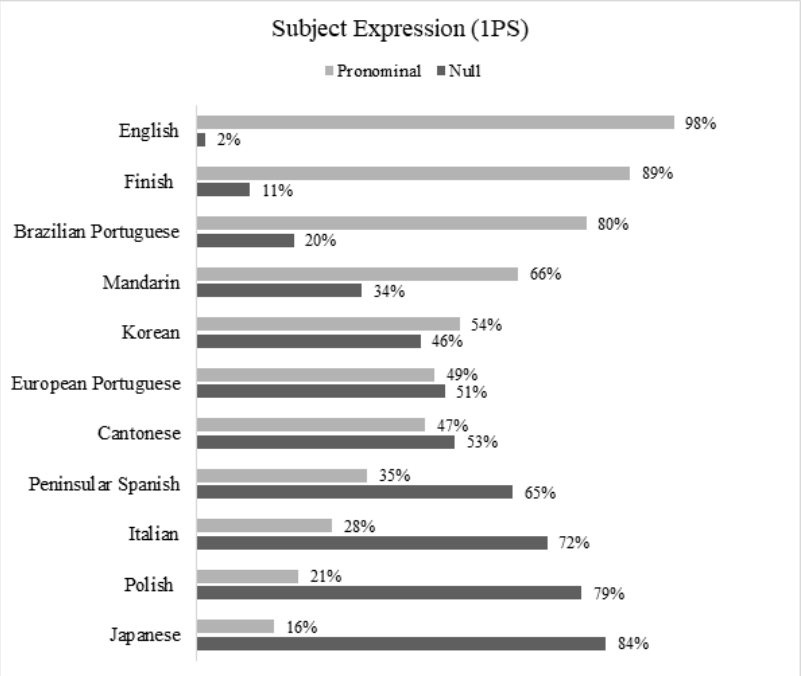
c. *Ø* estoy cantando

(2) a. *The child* was studying.

b. *I* am singing.

Graph 1 (below) observes this variable phenomenon across very diverse languages such as Japanese, Italian, Finnish, and English.

Graph 1. Null subject variation in 1PS in typologically different languages.



Source: Torres Cacoullos & Travis (2019), whose work is based on the following studies: for Japanese 84% (Lee & Yonezawa, 2008: 738, N=1571), Polish 79% (Chociej, 2011: 52, N=536), Italian 72% (Nagy p.c. cf. Nagy et al. 2011, N=224), Peninsular Spanish 65% (Posio, 2013: 269, N=787), Cantonese 53% (Nagy p.c. cf. Nagy et al. 2011, N=362), European Portuguese 51% (Posio, 2013: 269, N=704), Korean 46% (Oh, 2007: 466, N=433), Mandarin 34% (Jia & Bayley, 2002: 13, N=393), Finnish 11% (Helasvuo, 2014: 68, N=1793), English 2% (Torres Cacoullos & Travis, 2014: 22, N=6,600 (estimated)), and Duarte (2020), for two varieties of Portuguese (European Portuguese, 51% and Brazilian Portuguese, 20%).

Within the methodological framework employed in the World Atlas of Language Structures (WALS)[2] project (Dryer & Haspelmath, 2013), Dryer (2013) establishes six types of languages in accordance with the properties that they present in terms of subject expression in declarative sentences (Table 1).

Table 1. Subject expression and language typology (Dryer, 2013).

	Features / Properties	**N**	**Language**
I	Obligatory pronouns in the position of subject1	82	English, French
II	Subjects expressed by verb affix	437	Turkish, Spanish, Italian,
III	Subjects expressed by clitics on variable host	32	Chemehuevi
IV	Subject pronouns that occur in a different position from that of the subject	67	Longgu
V	Pronouns in the subject position but often omitted.	61	Japanese
VI	More than one trend without a dominant pattern	32	Finnish
	Total	711	

Source: Dryer (2013)

Type I languages, such as English in example (2) above, and French, are characterised by the realisation of the pronoun in the position of the subject. Type II, which includes European Portuguese, is the most frequently occurring paradigm. More than half of the 711 languages examined by Dryer fall into this category and have as a common feature the fact that the subject marker can be realised through verb affixes. The majority of romance languages are distributed across these first two types; most of them follow the dominant trend of this typology and fall into type II, but French is a type I language (Table 2).

Table 2. Typological classification of romance languages based on Dryer (2013).

Romance languages	**Pronoun**	**Typological classification (Dryer, 2013)**
European Portuguese	*(ele) come uma maçã*	Type II
European Spanish	*(el) come una manzana*	
Italian	*(lui) mangia una mela*	
Romanian	*(el) mananca un mar*	
Catalan	*(es) menja una poma*	
French	**(il) mange une pomme*	Type I
Romansh	**(Jeu) magliel in meil.*	

Source: Kaiser & Hack, 2010: 83

According to Kaiser & Hack (2010: 84), within the romance branch a third category exists in addition to these two types that should form part of type VI (Table 1), and for which the obligatory nature of the subject pronoun is conditioned by its morphological category of PERSON.

2 WALS project: https://wals.info/

> Toutes les études sur les dialectes septentrionaux de l'Italie s'accordent sur le fait que l'emploi du pronom sujet (clitique) dépend de la personne grammaticale, c.-à-d. qu'il est obligatoire ou quasi-obligatoire avec certaines personnes grammaticales, tandis que d'autres personnes présentent une variation libre par rapport à la présence du pronom (Heap 2000). La même conclusion peut être tirée par rapport à certaines variétés de l'occitan (SIBILLE 2007) ou du francoprovençal (DE CROUSAZ e SHLONSKY 2003; DIÉMOZ 2007; HINZELIN and KAISER 2008)[3].

For these authors, Romansh should not be included in Type I because it presents grammatical properties that distinguish it from French as much as from the other romance languages in Table 1. Unlike French, Romansh is a language with a rich verbal morphology. On the other hand, although the pronominal realisation of the subject is obligatory, it does not obey the same conditions as in French, for its omission is possible in contexts where it is post-verbal and in agreement with the grammatical PERSON. Such omission is very frequent in the second person, which makes Romansh an unusual language within the romance family, because it can also be included within type VI of Dryer's (2013) Typology.

> C'est l'omission du pronom sujet de la deuxième personne du singulier qui constitue la différence cruciale entre le romanche et les variétés romanes à *pro-drop* partiel: alors que l'omission de ce pronom est extrêmement fréquente en romanche, sa réalisation est obligatoire dans les autres variétés romanes – et encore plus obligatoire que celle de tout autre pronom sujet (KAISER & HACK, 2010: 90)[4]

This variable phenomenon has been widely studied using the theoretical model of Principles and Parameters of generative grammar, and it is within this framework that it has been interpreted as a Null Subject Parameter (RIZZI, 1982: 117-184; JAEGGLI & SAFIR, 1989: 1-44). From this perspective, a language such as Italian has the property [+Null Subject], with the indicators of a *pro-drop* language; it permits the omission of the subject therefore it is a Null-Subject Language (NSL). French is distinct from the parameter of Italian in that it is a non-Null-Subject Language (non-NSL) or [-Null Subject]. In contrast with Italian, French does not allow the omission of the subject. Nevertheless, and as Rizzi (1986) points out, it is not possible to account for interlinguistic variation only by establishing an opposition between

3 All studies of northern Italian dialects agree that the use of the subject pronoun (clitic) depends on the grammatical person, i.e., that it is obligatory or quasi-obligatory with certain grammatical persons, while other persons present free variation with respect to the presence of the pronoun (Heap, 2000: 66). The same conclusion can be drawn with respect to certain varieties of Occitan (Sibille, 2007) or Franco-provençal (De Crousaz & Shlonsky, 2003; Diémoz, 2007; Hinzelin & Kaiser, 2008).

4 It is the omission of the second-person singular subject pronoun that constitutes the crucial difference between Romansh and the partial *pro-drop* Romance varieties: while the omission of this pronoun is extremely frequent in Romansh, its realisation is obligatory in other romance varieties – and even more obligatory than that of any other subject pronoun (Kaiser & Hack, 2010: 90).

the obligatory vs. possible expression of the subject. As outlined above, other grammatical properties such as the grammatical category of person/number in the verb and/or the inversion of the position of the subject can have an influence, for example impeding the categorisation of a language as a NSL. In addition to needing to observe the verbal morphology, whether rich or not, of a given language, it is also necessary to take into account the different types of subject and their correlation with the alternation between null subject and pronominal subject. As such, it is necessary to distinguish between different types of subject:

(3) a. **null-argument, expletive subject with no referential content** (*il semble que le temps va changer*)

b. **argument subject, but one that is arbitrary or has an indefinite interpretation** (*On frappe à la porte*)

c. **argument subject with a definite reference** (*j'ai lu tout le livre*).

While the pronominal subject is obligatory in all contexts in French (cf. types of subjects) as illustrated in (1), the same cannot be said of standard European Portuguese. In this variety of Portuguese, the realisation of the null subject is obligatory in most contexts / for most types of subjects, although it is optional for argument subjects with a defined reference. These three types of subject may be *pro* or null, depending on the language, and they are characterised by a range of properties (Table 3).

Table 3. Types of subjects

Interpretation (types) of *pro*	PERSON	NUMBER
pro expletive	+	+
pro quasi-argument	-	+
pro expletive	-	-

Source: Sheehan (2016: 333)

Holmberg (2010) discusses a set of properties that interfere in the demarcation of the null subject parameter and proposes five types of languages (Table 4): in addition to the consistent or classical NSLs (Portuguese, Italian, Spanish) and the non-NSLs such as French and English that are described above, it is also important to consider partial NSLs (such as Finnish), expletive NSLs (German), and radical NSLs (Chinese, Japanese).

Table 4. Types of subjects and Homberg's (2010) typology

Types of Subjects	**Italian, European Portuguese, Greek**	**English, French**	**Finnish**	**German, Dutch**	**Chinese, Japanese**
null-argument subject, expletive, with no referential content	Not expressed	Expressed	Not expressed	Not expressed	Not expressed
quasi-argument subject, arbitrary or of indeterminate interpretation	Not expressed	Expressed	Variable	Expressed	Not expressed
Argument subject with definite reference	Variable	Expressed	Variable	Expressed	Not expressed
Holmberg's (2010) typology	**I Consistent null-subject languages**	**II Non-null-subject languages**	**III Partial null-subject languages**	**IV Expletive null-subject languages**	**V Radical null-subject languages**

Sources: Lobo (2016); Holmberg (2010); Figueiredo Silva (2017).

Another question that has attracted the attention of researchers is related to the different interpretations that can convey null subjects and overt subjects in NSLs (MONTALBETTI, 1984). These interpretations are related to pragmatic-discursive properties (CARDINALETTI & STARKE, 1999) and gave rise to the *Antecedent Position Hypothesis* (CARMINATI, 2002). In line with this hypothesis, Madeira, Fiéis & Teixeira (2021) observe that in European Portuguese, null subjects tend to refer to antecedents in the position of the subject, as in (4a.), while overt subjects tend to recover antecedents in other syntactic positions or extralinguistic referents, as in (4b.).

(4) a. A menina$_i$ beijou a avó$_j$ quando ***pro***[5]$_i$ chegou a casa

b. A menina$_i$ beijou a avó$_j$ quando **ela**$_{j/k}$ chegou a casa

Madeira, Fiéis & Teixeira (2021: 156)

5 *Pro* is a term used within the field of generative grammar, and especially in *Lectures on Government and Binding* (Chomsky, 1981), that denotes an empty category to account for phrasal elements that are not uttered (that are not produced phonetically) but have an effect on syntax. In this case, *pro* takes the subject position and receives the nominative case. *Pro* also forms part of a grammatical parameter – the Null Subject Parameter – of this theory, which seeks, amongst other things, to account for commonalities and divergences across all languages. From this perspective, in null subject languages (or *pro-drop* languages) a positive value is attributed to the parameter; while the languages in which the parameter is given a negative value are those which do not allow the omission of the subject, in other words, obligatory pronominal subject languages.

These authors also point out that the interpretative tendency of pronominal subjects can be cancelled by pragmatic factors, as **illustrated** in (5), whereby the null subject refers not to the constituent subject, but to an extra-phrasal antecedent:

(5) O João$_i$ não viu o despertador. O Rui$_j$ pensa que **pro$_i$** vai chegar atrasado.

Lobo (2013: 2323)

From the perspective of the formal model of generative grammar, the theoretical treatment of the Null Subject Parameter presents at least two challenges, which are observed by D'Alessandro (2014) and discussed further by Figueiredo Silva (2017: 194): on the one hand, there is a need to respond to empirical counterexamples, and in particular those which are based on dialectal data from European NSLs (such as varieties of Italian), and/or non-European varieties of Spanish (CAMACHO, 2013, for Dominican Spanish, among others), and/or other non-European NSLs (GILLIGAN, 1987); on the other, there is a need to resolve the problem presented by theoretical counter-evidence that jeopardises "the very existence of an empty category of the type of *pro*" and "the difficulty of attributing a clear theoretical status" to the micro-parameter of inflectional "richness". Even so, D'Alessandro underlines some indisputable advances: primarily, the correlation between null subjects and the morphology of rich agreement in inflectional languages; and the correlation between referential null subjects and obligatory expletive null subjects in NSLs.

Contributions from empirical variationist studies

This phenomenon has been widely researched, and an extensive literature has been published within both the functionalist approach and the generative model, producing significant results in terms of the distribution of expressed subjects and null subjects in natural languages. Nonetheless, certain questions remain unresolved, as Ayres (2021) observes in the introduction to her doctoral thesis. One such question relates to the factors that favour the use of null subjects versus pronominal subjects. A large proportion of publications in the variationist field deal with varieties of Portuguese and Spanish, whether these are peninsular and other European varieties, or the geographical varieties located principally on the American continent. Other variationist studies have focused on heritage varieties spoken by emigrants living for example in Toronto (NAGY, 2015) and New York (OTHEGUY, ZENTELLA & LIVERT, 2007).

Within the variationist field, four recent studies provide a synthesis of the linguistic factors that have come to be considered those which condition the variable null subject (Table 5).

Table 5. Most widely studied linguistic factors in the variation of pronominal vs. null subject expression

Linguistic factors	**Variables**
Type of subject	• Null-argument subject, expletive, no referential content • Subject of a meteorological verb • Argument subject, but arbitrary or of an indefinite interpretation • Argument subject with definite reference
Type of clause	• Finite clause • Co-ordinate • Subordinate
Distance	• Adjacent • Non-adjacent
"priming" effect (effect of the antecedent)	• A null subject • a pronominal subject • a nominal subject
Verbal semantic content	• Mental activity or epistemics • Stative verbs • Dynamic verbs
Verbal morphology	• Simple past • Imperfect • Present
Person and number	• 1st-person singular • 2nd-person singular • 3rd-person singular • 1st-person plural • 2nd-person plural • 3rd-person plural
Animacy	• Animate • Inanimate
Coreferentiality	• With a subject • With an indirect object • With a direct object • With an oblique object

Sources: Bouchard (2018), Posio (2018), Torres Cacoullos & Travis (2019), Flores & Rinke (2020).

These factors include the **type of subject, grammatical category of person/number, coreferentiality, and the "priming effect"** or antecedent, all of which have been referred to above. However, there are also other factors related to the adjacency and processing of elements, to the verbal semantic and morphological content, and to the feature of animacy of the subject referent.

On extralinguistic and social factors

By contrast with the diverse range and relevance of structural and linguistic factors that have been tested in empirical (variationist) studies, factors of a social or classically extralinguistic nature (age, sex, level of education) seem to have a lesser incidence on the variations (Table 6).

Table 6. Most widely studied sociolinguistic factors for variations in pronominal subject / null subject expression

Varieties of Portuguese	• *Level of formal education:* Bouchard (2018:15), São Tomé Portuguese (use of null subjects: 0.55 primary school; 0.50 high school; 0.45 university) • *Gender and living away from the community:* Lucchesi (2009: 181), Afro-Brazilian Portuguese • none: Duarte & Reis (2018: 187, PB (36 CORPORAPORT); Reis (2020) Madeira Island European Portuguese (Funchal, CORPORAPORT)
Varieties of Spanish	• none: Otheguy et al. (2007:778) Spanish-speaking residents of New York. • *age:* Orozco & Guy (2008): E. Barranquilla, Colombia; Alfaraz (2015): Santo Domingo, Dominican Republic; Lastra & Butragueño (2015): Mexico City (Mexico); Prada Pérez (2015), bilingual Catalan-Spanish speakers, Menorca, Spain. • *gender:* (Posio 2016)

Sources: Bouchard (2018), Lucchesi (2009), Duarte & Reis (2018), Reis (2020), Othegui et al. (2007), Orozco &Guy (2008), Alfarraz (2015), Lastra & Butragueño (2015), Prada Pérez (2015), Posio (2016).

In some studies, just one of the variables has a statistical effect – such is the case with the *age* variable in varieties of Spanish (OROZCO & GUY, 2008; LASTRA & BUTRAGUEÑO, 2015; ALFARAZ, 2015), or the *sex* variable in varieties of peninsular Spanish and peninsular European Portuguese (POSIO, 2016). In other studies, however, none of the three variables is shown to be statistically relevant, as will be discussed in more detail later in this article.

The production of null subjects in spoken varieties of Portuguese

In its traditional form of standard European Portuguese, Portuguese as a consistent Null Subject Language is marked by the possibility of the omission of the referential subject. In addition to a morphology of rich verbal agreement in which the verbal form already possesses affixes of Person/Number, standard European Portuguese exhibits a set of properties (LOBO, 2016: 562) that are illustrated by the examples given in (6) to (9):

(6) Possibility of referential null subjects

a. *Ø* Fomos à universidade.

b. Ø Conheço a Maria.

(7) Possibility of expletive null subjects

c. Choveu.

d. Parece que vai chover.

(8) Possibility of inversion of the order of subject-verb

e. Nasceu o meu sobrinho!

f. Floriram as cerejeiras em Trás-os Montes!

(9) Possibility of extraction of the subject at a distance

g. Quem você/ocê disse que vai via

h. Quem você/ocê acha que foi o vencedor

Semantic factors such as *animacy* must also be considered, because they take on great importance in varieties of Portuguese. In standard European Portuguese, there is a tendency for overt pronominal subjects to be [+animate] and for pronominal subjects that refer to [-animate] antecedents to be null (BARBOSA, DUARTE & KATO, 2005; LOBO, 2016):

(10) a. O João leu que o *tsunami de 2004*$_i$ foi o desastre natural mais mortal da história. No total, *#ele*$_i$ tirou a vida a 230 000 pessoas.

b. O João leu que *Jack o estripador*$_i$ é um dos criminosos mais famosos de sempre. No total, *ele*$_i$ tirou a vida a 5 pessoas.

Not all varieties of Portuguese observe the same pattern of properties. The literature on the differences between European Portuguese and Brazilian Portuguese is extensive. Spoken Brazilian Portuguese avoids leaving the position before the verb empty (KATO, 2002), as the following examples illustrate:

(11) a. *Ø* Durmo cedo.
a'. *Eu* **durmo cedo.**
b. *Ø Conserta sapato(s)*
b'. *Aqui* conserta sapato.
c. *Ø* Vira à esquerda.
c'. *Cê* vira à esquerda.
d. *Ø* Dormem ali os meninos.
d'. *Ali* dormem os meninos
e. *Ø* Chove em São Paulo.
e'. *São Paulo* chove.
f. *Ø* Furou o pneu do Hamilton.
f'. *O Hamilton* furou o pneu.
g. *Ø* Xinguei o cara.
g. *Daí* xinguei o cara.

Table 7 shows some of the differences between European Portuguese and Brazilian Portuguese in relation to the possibility of the production of null subjects, considering the correlation of the subject with other properties such as the type of subject, coreferentiality, and the animacy or not of the subject, as well as other grammatical possibilities within these two varieties, as is the case in the production of more complex structures such as that of double subjects.

Table 7. Principal structural differences in the production of null subjects (vs. overt subjects) in EP and BP

	European Portuguese	**Brazilian Portuguese**
Omission of referential subjects	Ø Fomos à praia	A gente foi à praia / Nós fomos à praia.
Arbitrary null subjects	**É** assim que *(se) faz o doce Nesse hotel não *(se) pode entrar na piscina bêbado	É assim que faz o doce Nesse hotel não pode entrar na piscina bêbado.
Coreferentiality	O chefe$_i$ disse ao amigo$_j$ que *pro*$_i$ precisava de descansar O chefe$_i$ disse ao amigo$_j$ que ele$_j$ precisava de descansar	A Ana$_i$ disse à Rosa$_j$ que ela$_{i/j}$ precisava de descansar O povo brasileiro$_i$ acha que ele$_i$ tem uma grave doença
Animacy	A casa$_i$ virou um filme quando teve de ir abaixo.	A casa$_i$ virou um filme quando ela$_i$ teve de ir abaixo.
Double subjects	*A Clarinha ela cozinha que é uma maravilha.	A Clarinha ela cozinha que é uma maravilha.
Arbitrary pronominal subjects	**Ø**expl Deviam ensinar o amor às crianças. Øexpl Chove muito nessas florestas.	*Eles* deviam ensinar amor às crianças. *Essas florestas* chovem muito. (Kato e Duarte, 2014)

Source: Lobo (2016: 572).

By contrast with European Portuguese, Brazilian Portuguese reveals a tendency to complete the subject, whether it is an argument or referential, quasi-argument, or arbitrary. In this variety, pronominal referential subjects may also correspond to [-animate] referents, meaning that they are not subject to the semantic restrictions observed in European Portuguese. On the other hand, and contrary to European Portuguese, the subject in the main clause and the pronominal subject in the subordinate clause may be coreferential in Brazilian Portuguese. The duplication of the subject by a pronoun is possible in Brazilian Portuguese and not in standard European Portuguese. In both varieties, expletive subjects always correspond to null subjects (*Ele parece que *ele vai chover amanhã). With regard to arbitrary subjects, in Brazilian Portuguese there is the use of the generic null, as Kato & Duarte (2014) underline, and the disappearance of the indefinite *se* that is used in European Portuguese (GALVES, 1987). This

variant competes with other that have an expressed pronominal subject such as *a gente* or *você*, as in the examples (12) given by Kato & Duarte (2014: 9):

(12) a. *Ø* não pode fumar aqui.
b. *A gente* não pode fumar aqui.
c. *Você* não pode fumar aqui.

For Holmberg *et al* (2009), this is a construction found in *partial* Null Subject Languages, but not in consistent Null Subject Languages such as Italian and Spanish, which require the clitic *se*.

It is possible therefore to observe that Brazilian Portuguese has significantly reduced the contexts in which the null subject may be realised. Cyrino, Duarte & Kato (2000) observe on the basis of their work on syntactic specificities in the area of the subject and the direct object in Brazilian Portuguese that referentiality is very important for the realisation of the pronoun. To account for these specificities, they propose the concept of a referential hierarchy, in (i):

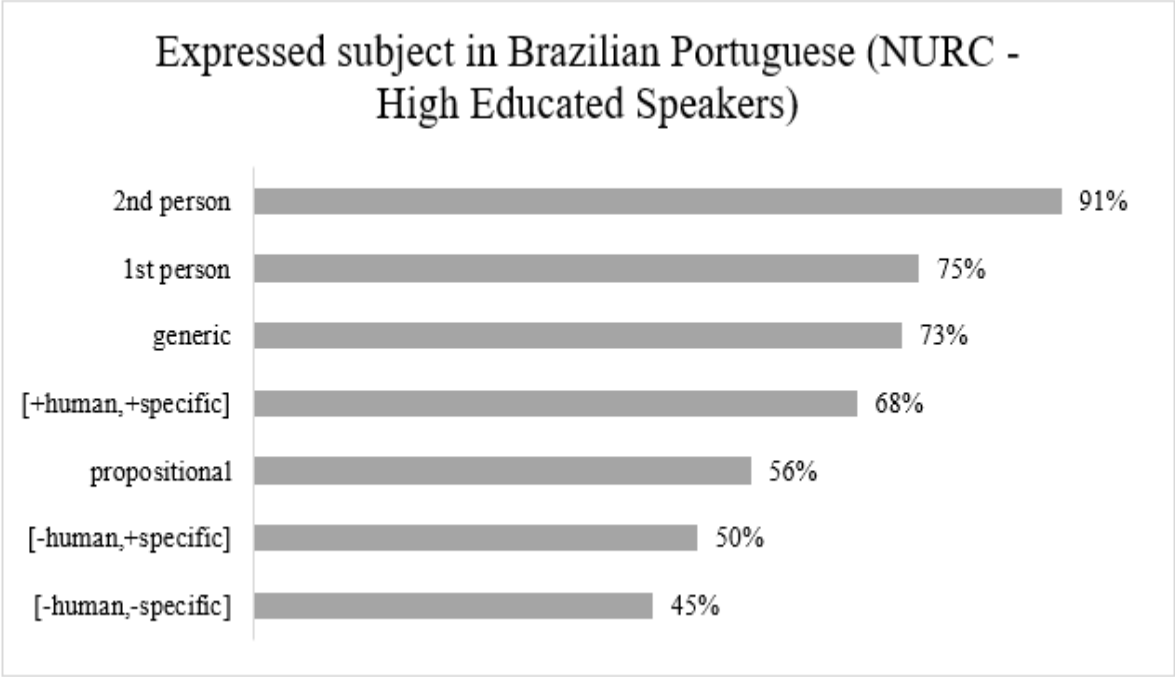
(i) Referential hierarchy (Cyrino, Duarte & Kato, 2000: 39)

null-argument	**proposition**	**P3**	**P2 P1**
		+human	±human
		±specific	

[-referential] < ----------X----------X----------> [+referential]

The pronominalisation of the subject is implemented when the conditions are in place for greater referentiality, in other words, in the 1st and 2nd persons, with [+human]; and it diminishes progressively with antecedents of lesser referentiality, first in the 3rd person [+human] followed by [-human]. Arbitrary [-referential] subjects offer greater resistance to pronominal omission and are therefore null, and in this they are consistent with the standard variety of European Portuguese. Graph 2, below, illustrates the way in which this referential hierarchy is applied in Brazilian Portuguese, using data from the Norma Urbana Oral Culta (NURC) of educated speech from Rio de Janeiro, first analysed by Duarte (1995) and referred to in Kato & Duarte (2014).

Graph 2. Expressed subjects (vs. null subjects) within the referential hierarchy of BP



Source: Kato & Duarte (2014: 5)

The differences between European Portuguese and Brazilian Portuguese in relation to the possibility of the production of null subjects are summarised here, and they lead to the conclusion that the grammar of the Brazilian variety of Portuguese has been the object of a process of change during which it has lost properties that are prototypical of a Null Subject Language such as European Portuguese. The hypothesis that BP can be considered a *partial* NSL, as proposed by Holmberg & Sheehan (2010) has not drawn a consensus. Kato (2020: 77) outlines four hypotheses for the direction of change underway in BP, which are that it will become: i) a non-Null Subject Language of the same type as French and English; ii) a *radical* Null Subject Language, like Chinese and Japanese; iii) a *partial* Null Subject Language such as Finnish; or iv) a "semi" Null Subject Language similar to Icelandic. Kato (2020: 89) reviews each of these hypotheses and ultimately argues for a framework in which a competition between grammars can be observed in Brazilian Portuguese – between the innovative, which is acquired as a core grammar, and the conservative, which is learned as a second grammar through formal instruction. According to the author, this core grammar would be similar to that of Icelandic, "com sujeitos referenciais que consistem em um paradigma de pronomes fracos livres como o inglês e o francês, mas com elementos nulos no lugar de sujeitos não referenciais" ["with referential subjects that consist in a paradigm of free weak pronouns like in English and French, but with null elements in the place of non-referential subjects"] (KATO, 2020: 89).

The *corpora*

The comparison of results is based on *corpora* comprising the findings of empirical research detailed in fourteen publications, listed in Table 8 below.

Table 8. Corpora of publications on the production of null subjects in spoken varieties of Portuguese

Varieties	Studies	Corpora
Afro-Brazilian Portuguese	Lucchesi, 2009	Corpus Vertentes
Dialectal Brazilian Portuguese	Myamoto & Kailer, 2019	Sample of the Atlas Linguístico do Brasil (ALiB)
Brazilian Portuguese (Rio de Janeiro)	Duarte,1995; Soares, 2017;	NURC (Rio de Janeiro); Corpus Concordância[6] - Rio de Janeiro
European Portuguese (Lisbon)	Duarte & Reis, 2018; Duarte, 2020	Corpus Concordância (Lisbon)
Dialectal European Portuguese	Carvalho, 2009	Corpus Cordial-SIN
Madeira Island European Portuguese	Reis, 2020	*Corpus* Concordância (Funchal)
São Tomé Portuguese	Bouchard, 2018	*Corpus* Bouchard (L1, Bilingual)
Angolan Portuguese	Oliveira & Santos, 2007	*Corpus* Chavagne (2005)
Popular Angolan Portuguese	Teixeira, 2012	*Corpus* Teixeira (L1, Bilingual)
Mozambican Portuguese	Oliveira, 2016	*Corpus* Anna Jon-And (L2, Maputo)
Heritage Portuguese (Hamburg) (HP Hamburg)	Flores & Rinke, 2020	*Corpus* of Heritage Portuguese (1st and 2nd Generations - Hamburg)

Source: Elaborated by the author

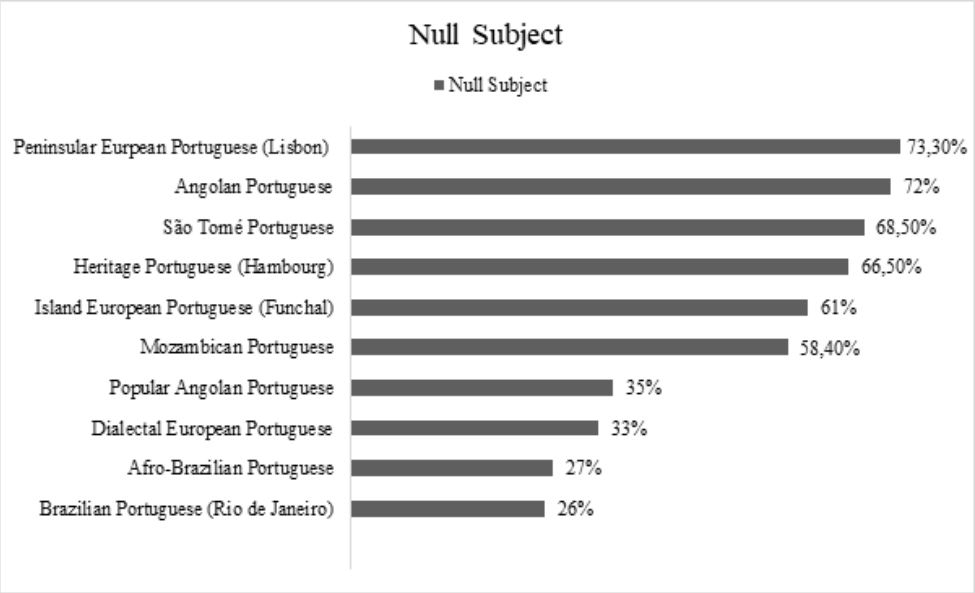
The selection seeks to bring together contributions that are focused on data related to spoken varieties of Portuguese. It covers geographical and social varieties, and different modes of acquisition of Portuguese (African, Brazilian, and European varieties, rural and urban varieties, L1 and also L2 with or without bilingualism, Heritage Languages). Research focused on data pertaining to written production was not included in this study, and discussions that centre almost exclusively on the theoretical discussion of the null subject parameter within the framework advocated by generative grammar were also excluded.

6 The Concordância *Corpus* currently forms part of the COPORAPORT project.

Results

The global results (Graph 3) show the configuration of a *continuum* on which, as expected, the peninsular variety of European Portuguese is located at one extreme that is characterised by a higher percentage of null subjects; at the opposite end, various varieties of Brazilian Portuguese are found, in all of which the percentage of null subjects is much smaller. Between these two broad varieties of European Portuguese and Brazilian Portuguese are located the Madeira Island variety of European Portuguese and various African varieties of Portuguese.

Graph 3. Variations in Null Subject expression in spoken varieties of Portuguese.

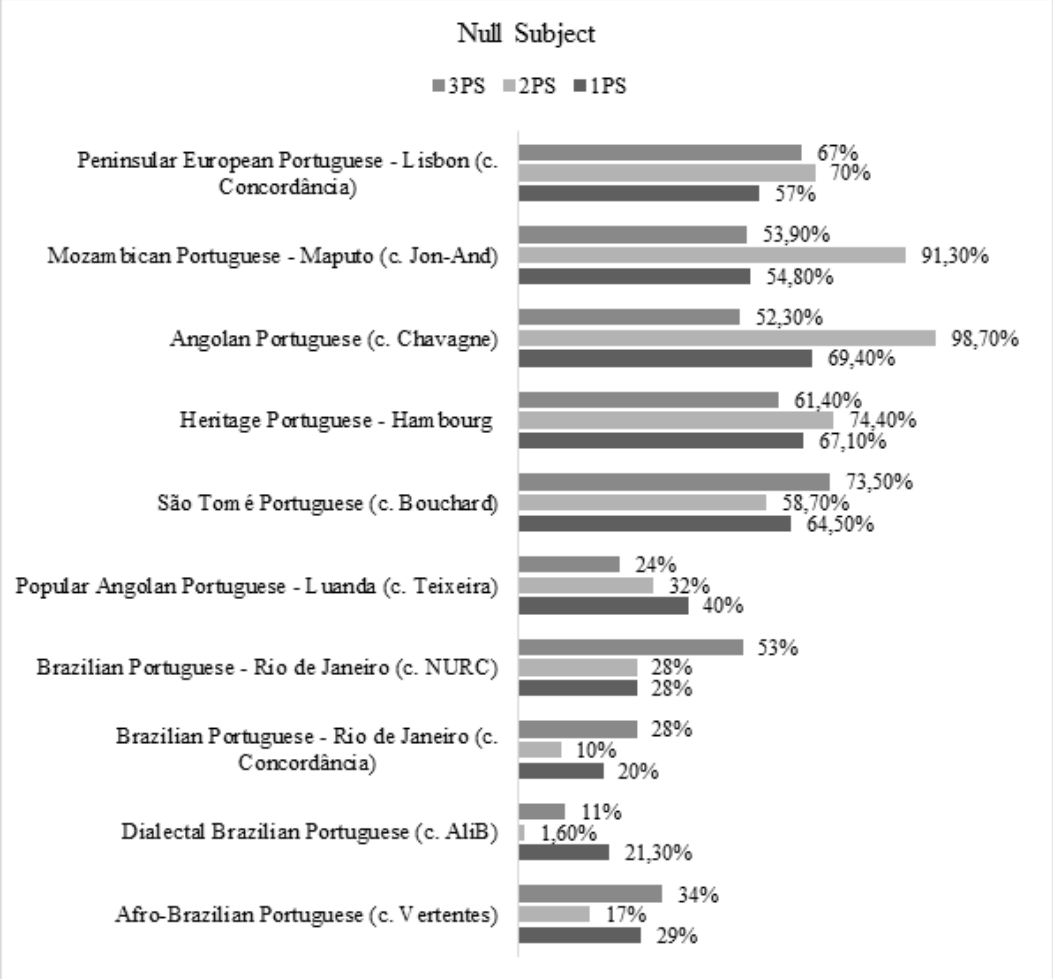


Sources: Peninsular European Portuguese (Duarte, 2000), Madeira Island European Portuguese (Reis, 2020); Dialectal European Portuguese, based on CORDIAL-SIN (Carvalho, 2009); Brazilian Portuguese (Duarte, 2000); Afro-Brazilian Portuguese (Luccchesi, 2009); São Tomé Portuguese (Bouchard, 2018), Angolan Portuguese (Oliveira & Santos, 2007), Popular Angolan Portuguese (Teixeira, 2012), Mozambican Portuguese (Oliveira, 2016); Heritage Portuguese (Germany) (Flores & Rinke, 2020).

Contrary to what might be expected, the results pertaining to the production of null subjects in the socially opposed varieties of Afro-Brazilian Portuguese (LUCCHESI, 2009) and educated Brazilian Portuguese (with a sample from the NURC, DUARTE, 1995; SOARES, 2017) give very similar outcomes of 27% and 26% respectively. Therefore, the change that has happened in BP must not be related to the level of formal education of the speakers, nor is their background related to language contact, as Lucchesi (2009: 176) observes. For Lucchesi, this situation of convergence arises from two developmental trajectories: that of the educated variety and the structural reorganisation that is provoked by the substitution of the pronouns *tu* and *nós* by *você* and *a gente*; and that of popular varieties, in which inflected morphemes of person and number are dropped in the irregular linguistic transmission that occurs through the contact of Portuguese with African languages.

The results of the uses of the null subject distributed across the 1PS, 2PS, and 3PS are presented in Graph 4 below:

Graph 4: Variations in Null Subject expression by PERSON/NUMBER (1PS, 2PS, and 3PS) in spoken varieties of Portuguese.



Sources: Afro-Brazilian Portuguese (*corpus* Vertentes, cf. Luccchesi, 2004); Dialectal Brazilian Portuguese (*corpus* ALiB, Miyomoto & Kailer, 2019), Brazilian Portuguese – Educated Speakers (*corpus* NURC, Soares, 2017) Brazilian Portuguese (*corpus* Concordância, Duarte, 2020), Peninsular European Portuguese (*corpus* Concordância, Duarte, 2020), São Tomé Portuguese (Bouchard, 2018), Popular Angolan Portuguse (Teixeira, 2012), Mozambican Portuguese (*corpus* Jon-And, Oliveira, 2016) and Heritage Portuguese - Hamburg (Flores & Rinke, 2020).

There is a clear tendency towards the more frequent omission of the subject in the context of the 3PS in all varieties, although less so in the Popular Angolan variety (Luanda) and in the BP variety in the ALiB data. In these varieties, null subjects occur more frequently in the 1PS.

In relation to the other linguistic varieties, and as Bouchard (2018) observes, the studies selected do not present results that can be usefully compared because they make use of different variables and analytical methods. Some exceptions, however, are the recent studies by Duarte (2020) and Reis (2020) based on the Concordância *corpus*, which includes samples of Brazilian

Portuguese (Rio de Janeiro: Nova Iguaçu and Copacabana), a sample of Peninsular European Portuguese (Lisbon: Cacém and Oeiras), and a sample of Madeira Island European Portuguese (Funchal). The results presented in Chart 1 below account for the correlation between *animacy* and the production of null subjects.

Chart 1. Variations in Null Subject expression in the 3PS and its correlation with the semantic features of the Subject in spoken varieties of European Portuguese and Brazilian Portuguese.

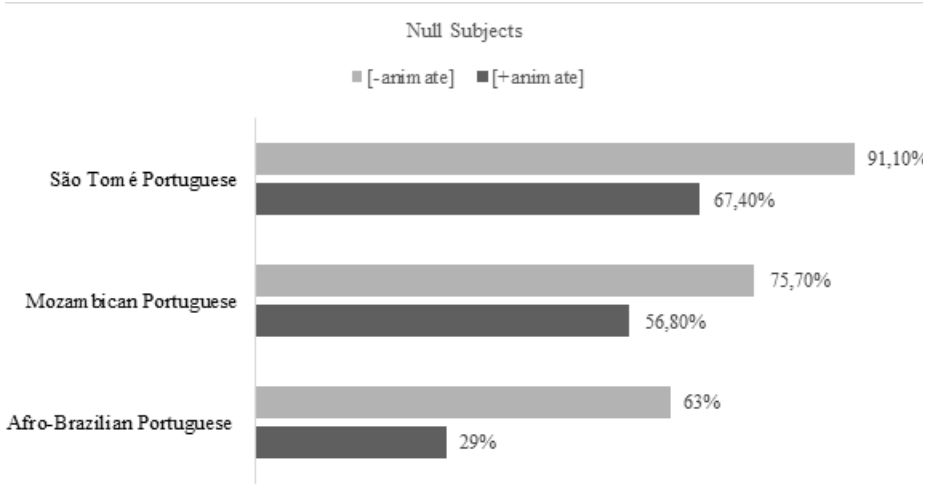
	Peninsular European Portuguese input 0.756			**Madeira Island European Portuguese (Funchal) input 0.748**			**Brazilian Portuguese (Rio de Janeiro) input 0.248**		
Features	**N/T**	**%**	**P.R.**	**N/T**	**%**	**P.R.**	**N/T**	**%**	**P.R.**
[-animate /-specific]	12/12	100%	-	12/12	100%	-	7/12	58%	**0.863**
[-animate /+specific]	137/142	96.5%	**0.942**	180/187	96%	**0.919**	73/173	42%	**0.692**
[+animate /-specific]	191/246	78%	**0.562**	155/194	80%	**0.547**	62/191	32.5%	**0.555**
[+animate /+specific]	308/559	55%	**0.307**	173/403	43%	**0.229**	189/803	23.5%	**0.437**
	range		*0.635*			*0.690*	*range*		*0.426*
	Log likelihood = -440.725 Significance = 0.000			Log likelihood = -350.925 Significance = 0.000			Log likelihood = -624.928 Significance = 0.000		

Sources: Peninsular European Portuguese and Brazilian Portuguese (Duarte, 2020) and Madeira Island European Portuguese (Reis, 2020)

It can be observed in these results that the referential subjects [-animate, +specific] and [-animate, -specific] in the two varieties of European Portuguese correspond to categorical (100% in both) and almost categorical (96.5% and 96%) null subjects, while this does not occur in Brazilian Portuguese. In this latter variety, a significant tendency towards the use of null subjects in these contexts (58% and 42%) can also be seen, although in a much smaller number. The most striking contrast is evident in contexts where the feature [+animate] is predominant: BP tends to use overt subjects (67.5% and 76.5%), while European Portuguese uses null subjects (78% and 55%). Contexts with the feature [-animate] offer the greatest resistance to the implementation of the forms of the subject expressed in Rio de Janeiro varieties of Brazilian Portuguese.

It is possible to compare in absolute terms the effect of animacy on the Afro-Brazilian variety and on two African varieties of Portuguese (Mozambican Portuguese and S**ão** Tomé Portuguese) in other studies.

Graph 5. Variations in Null Subject expression according to the variable animacy of the subject in three spoken varieties of Portuguese.



Sources: Afro-Brazilian Portuguese (Lucchesi 2009); Mozambican Portuguese (Oliveira, 2016, with the Jon-And corpus of Maputo, L2); São Tomé Portuguese (Bouchard, 2018).

The same pattern occurs across the three varieties: the feature [-animate] stimulates considerably the production of null subjects. However, only the São Tomé variety seems to come close to the European Portuguese data.

The contrasts between varieties of European Portuguese and Brazilian Portuguese may be less marked when we consider the data for Dialectal European Portuguese. Carvalho's (2009) work has shown that the percentage of use of *null subjects* is not uniform in EP and that it produces figures that do not correspond to a classical or consistent Null Subject Language. In her study of dialectal (peninsular and island) varieties of EP, based on a sample from CORDIAL-SIN, the author presents percentages for pronominal subjects that are higher than those for null subjects. Only one locality does not conform to this pattern: Figueiró, in the district of Guarda. In this locality, the percentage for null subjects (64%) is superior to that for pronominal subjects (36%) (Chart 2).

Chart 2. [+specific] subjects in dialectal European Portuguese (data from CORDIAL-SIN)

Locality	Pronominal subject	Null subject	Total
Alcochete	48 (69%)	22 (31%)	70
Alvor	78 (76%)	25 (24%)	103
Cabeço de Vide	326 (92%)	29 (8%)	355
Camacha	85 (62%)	51 (38%)	136
Ponta da Garça	288 (93%)	22 (7%)	310
Serpa	19 (53%)	17 (47%)	36
Arcos	167 (65%)	90 (35%)	257
Outeiro	86 (55%)	70 (45%)	156
Figueiró	70 (36%)	123 (64%)	193
Covo	561(67%)	276 (33%)	837
	1728 (70%)	725 (30%)	2453

Source: Carvalho (2009: 134)

The most surprising result corresponds to the locality of Ponta da Garça on the island of São Miguel (Azores), with 93% of expressed subjects – a higher percentage than that for the BP variety that has the highest percentage for completion of the subject in the publications selected for this study (73%).

In relation to the effect of the grammatical person/number variable on the use of null subjects, the dialectal data for EP reveal a certain discrepancy with regard to those expressed in the varieties of EP that are included in the Concordância *Corpus*, as attested in Chart 3.

Chart 3. PERSON/NUMBER variable and subject expression in dialectal European Portuguese data from CORDIAL-SIN

	1PS (eu)	**2PS (tu)**	**3PS (ela)**	**3S/1P (a gente)**	**1PP (nós)**	**3PP (eles)**	**Total**
Null subjects	417 27%	20 22%	209 42%	10 11%	20 37%	46 57%	**721** **30%**
Overt subjects	1132 73%	71 78%	296 58%	83 89%	34 63%	35 43%	**1651** **70%**
	1548	91	505	93	54	81	**2372**

Source: Carvalho (2009: 145)

Subject expression produces very high values in the 1PS (73%) and in the 2PS (78%). This tendency is counter to that shown in Duarte's (2020) results, which are more in accordance with the properties of a Null Subject Language, with percentages for subject expression of 43% and 30% respectively.

Another result worthy of mention here comes from an island variety of EP, that also contrasts with the properties attributed to urban varieties of European Portuguese and Brazilian Portuguese. A study by Casalta (2017), based on data taken from the locality of Camacha on the island of Porto Santo (Madeira Archipelago) and made available in the CORDIAL-SIN *corpus* shows that in 14 arbitrary subjects there are 10 pronominal subjects. The examples at (13) are interpreted as arbitrary pronominal subjects [+animate] because they represent the indefinite referent *people*.

(13) a. ***eles*** *agora* ***costumam*** *de lavar a vinha.* (PST02)

b. *Aqui,* ***eles andaram*** *fazendo e deram com água.* (PST07)

These examples diverge from the standard variety of European Portuguese. In that variety, arbitrary subjects in the 3PP are categorically null and an expressed pronoun in the 3PP can only have a specific referential interpretation.

Finally, the construction of the double subject should also be mentioned. This is a complex structure that combines the pronoun *a gente* with the clitic *se*. It is very common in Madeira Island European Portuguese, whether in the urban variety (Funchal) or in the rural varieties (HENRIQUES, 2019), and it is also found in peninsular varieties of European Portuguese. The

double subject expressed by *a gente se* has an arbitrary interpretation and a complex nature, because its referential reach can oscillate between the generic [-specific, +impersonal] and the deictic [+specific, -impersonal], as in the following examples:

(14) *generic:* Havia um arraial ali que **a gente** chama**-se** "Maio".
definite: Bordava-**se** o sábado, só o domingo é que ia-**se** p'o caminho brincar.
specific: **A gente** era-**se** garotos pequenos.
deictic: Ah, pega-me às costas. **A gente** chega-**se** lá num instante.

(HENRIQUES, 2019: 57-58)

This construction is found in non-standard varieties of European Portuguese and is more frequent in the variety spoken on island of Madeira. It constitutes a variant of the double subject construction. It is rooted in European Portuguese because it maintains the clitic *se*, contrary to what is observed in the varieties of Brazilian Portuguese:

(15) a. No Brasil, não se usa mais saia (European Portuguese)

b. No Brasil, não Ø usa mais saia. (Brazilian Portuguese)

As Barbosa, Duarte & Kato (2005) recognise, in Brazilian Portuguese other variants of the overt subject are used, expressed by *a gente* and *você*, as shown in (16):

(16) a. E se *a gente* falar que não tem? (KATO & TARALLO, 1986: 347)
b. *Você, no Canadá, você pode ser o que você quiser.* (BARBOSA, DUARTE & KATO, 2005: 19)

Furthermore, in Mozambican Portuguese spoken in Maputo, it is possible to observe indeterminate subjects that are similar to those in the Brazilian Portuguese data: using the third-person verb without *se* and without any special marker (17), and with the use of the pronoun *você* (18):

(17) a. Às vezes há dias que não *Ø* consegue nada. (MZ17)
b. Aqui agora já não *Ø* faz mais isso. (MZ4)

(18) a. Só quando *você* tem aquela classe pequena, mas nessa classe pequena *você* aproveitava. (MZ12)
b. Quando *você* vai para a escola, *você* não vai aprender nada. (MZ12)
(Oliveira, 2016: 67)

Social variables have not merited much attention in existing research. In many works, information pertaining to these variables is not made available. The *education level* variable was found to be statistically relevant in the work of Bouchard (2017) on the African variety of STP; the *age* variable in the MP variety (OLIVEIRA, 2016), and *sex* in the sample of Afro-Brazilian Portuguese analysed by Lucchesi (2009).

Table 9. Social variables and the production of null subjects in spoken varieties of Portuguese.

	age	**sex**	**level of education**	**lived outside the community**	**contact with other languages**
Afro-Brazilian Portuguese	No	Male 0.52	X	Never 0.53	X
Dialectal Brazilian Portuguese (*Corpus* ALiB)	No	No	No	X	X
Brazilian Portuguese (*Corpus* Concordância Rio Janeiro)	No	No	No	X	X
European Portuguese (*Corpus* Concordância Lisbon)	No	No	No	X	X
Dialectal European Portuguese (CORDIAL-SIN) *	X	X	X	X	X
São Tomé Portuguese (*Corpus* Bouchard)	-	-	Primary 0.55 Secondary 0.50	-	**
Angolan Portuguese (*Corpus* Chavaigne)*	X	X	X	X	X
Popular Angolan Portuguese *Corpus* Teixeira)	X	X	X	X	X
Mozambican Portuguese (*Corpus* Jon-And)	35-59 0.55				
Heritage Portuguese (*Corpus* Hamburg, Germany) ***	X	X	X	X	X

* Profile of informants: illiterate, elderly, born and always lived in the same locality

** Level of education variable

*** Forms a single extralinguistic variable: 1st generation vs 2nd-generation speakers

X information unavailable/not published.

Sources: Afro-Brazilian Portuguese (Lucchesi, 2009); Mozambican Portuguese (Oliveira, 2016, with the Jon-And corpus (Maputo, L2); São Tomé Portuguese (Bouchard, 2018); Dialectal European Portuguese (Carvalho, 2009, with data from CORDIAL-SIN); Brazilian Portuguese (Duarte, 2020, Corpus "concordância-RJ"); Dialectal Brazilian Portugueses (Miyomoto & Kailer, 2019, with data from) ALiB (Atlas Linguístico do Brasil); Angolan Portuguese (Oliveira & Santos, 2007 – corpus Chavaigne); Popular Angolan Portuguese (Teixeira, 2012); Heritage Portuguese – Hamburg (Flores & Rinke, 2020).

Final Considerations

Following this examination of the production of null subjects in spoken varieties of Portuguese, the difficulties in making a comparison of the data pertaining to these varieties are patently clear. The research methodology used in each of the different studies selected and analysed here is variable, and this makes the task of comparing the results difficult and at times impossible.

It has only been possible to compare the different varieties according to the total percentages of null subjects, and their distribution across the grammatical category of Person (1st, 2nd, 3rd). In these two comparisons, a gradual increase in the percentages of null subjects is observed in the direction of the varieties of European Portuguese spoken in Lisbon (Oeiras and Cacém), when traced from the starting point of the Afro-Brazilian variety and other varieties of Brazilian Portuguese at the opposite end of the scale. The hypothesis that arises from this exercise is that this sphere is barely permeable to the effects resulting from contact with other languages. Whether in the situation of the variety of Heritage Portuguese spoken in Hamburg, Germany, where the dominant language of German – an expletive Null Subject Language – is in contact with the consistent Null Subject Language of Portuguese, or in the situation of African varieties which have to a greater or lesser extent become native varieties of Portuguese, the percentages of null subjects are consistent with the typological determination of European Portuguese.

The analysis of other structural factors such as the correlation of the production of null subjects and the semantic feature of *animacy* has only been partially possible. This illustrates the principal difficulty faced in this study: the differences in methodological approaches outlined above and observed by Bouchard. As emphasised in the literature, the [-animate] feature is realised categorically by null subjects in varieties of European Portuguese and almost categorically in African varieties such as São Tomé Portuguese. In Brazilian Portuguese, this semantic feature presents the strongest resistance to change in the direction of overt subject expression.

In general, the low significance of social / extralinguistic factors should be pointed out. Indirectly, and although it does not seem to have an effect on the varieties of Brazilian Portuguese, the *education level* factor proves to have a greater impact on varieties of European Portuguese. In European Portuguese, dialectal varieties that are typically rural and produced by informants with a low level of formal education reveal general rates of null subjects far inferior to those of the social urban varieties of the Lisbon region, which may indicate the necessity for a more detailed examination of European Portuguese data as a complex object in which various structural conditions interact. Furthermore, the convergence in uses of null subjects across Afro-Brazilian varieties and more highly educated varieties of Brazilian Portuguese (Rio de Janeiro) seems pertinent and worthy of deeper reflection.

Finally, it is worth highlighting the role of dialectal diversity observed in the linguistic area of Portuguese, and especially in the sphere of European Portuguese, which occurs in similar fashion to what happens in other romance languages such as Italian. The dialectal data for European Portuguese provide empirical counter-evidence against the parameters and typologies that are based almost exclusively on the possibility or otherwise of the production of null subjects. In this field, it would perhaps be preferable to face the possibility of recourse to the concept of *gradation* in variations, and to include it conceptually within the framework of parametric change, as Oliviéri (2004: 119-120) and Hinzelin & Kaiser (2012: 259-260) advocate in respect of the structural variations observed in the romance family.

References

ALFARAZ, G. G. *Variation of Overt and Null Subject Pronouns in the Spanish of Santo Domingo*. Washington, D.C.: Georgetown University Press, p. 3-16, 2015.

ALiB. https://alib.ufba.br.

ARENDS, J. Simple grammars, complex languages, *Linguistic Typology*, 5, p.180-182, 2001.

ARIEL, M. *Accessing Noun-Phrase Antecedents*. London: Routledge, 1990.

AUGER, J. Les pronoms clitiques sujets en picard: une analyse au confluent de la phonologie, de la morphologie et de la syntaxe. *Journal of French Language Studies*, 13, p.1–22., 2003. https://doi.org/10.1017/S0959269503001066.

AYRES, M. R. *Contextos licenciadores de sujeitos nulos em português brasileiro*. PhD Thesis. Instituto de Letras da Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil, 2021.

BARBOSA P.; DUARTE, M.E.L.; KATO, M. Null subjects in European and Brazilian Portuguese. *Journal of Portuguese Linguistics*, 4, p. 11–52, 2005.

BENINCÀ, P. Subject clitics and particles in Provençal, *Probus,* 26(2), p. 183-215, 2014.

BOUCHARD, M-E. Subject Pronoun Expression in Santomean Portuguese. *Journal of Portuguese Linguistics*, 17 (5), pp. 1–29, 2018. DOI: https://doi.org/10.5334/jpl.19

CAMACHO, J. A. *Null Subjects*. Cambridge: Cambridge University Press, 2013.

CAMERON, R. *Pronominal and Null Subject Variation in Spanish: Constraints, Dialects, and Functional Compensation*. Unpublished Ph.D. thesis, University of Pennsylvania, Philadelphia, PA, USA, 1992.

CAMERON, R.; FLORES-FERRÁN, N. Perseverance of subject expression across regional dialects. *Spanish in Context*, 1, p.41–65, 2004. DOI: https://doi.org/10.1075/ sic.1.1.05cam.

CARDINALETTI, A. Subjects and clause structure. In: HAEGEMAN, L. (ed.) *The new comparative syntax*. London: Longman, 1997, p.33–63.

CARDINALETTI, A.; STARKE, M. The typology of structural deficiency: A case study of the three classes of pronouns. In: RIEMSDIJK, H. van (ed.), *Clitics in the languages of Europe*. New York: Mouton de Gruyter, 1999, p. 145-233.

CARMINATI, M.N. *The processing of Italian subject pronouns*, 2002. *Doctoral Dissertations Available from Proquest. AAI3039345. https://scholarworks.umass.edu/dissertations/AAI3039345*

CARVALHO, G.A. *A realização do sujeito na fala do portugues europeu*. PhD Thesis. Universidade Estadual Paulista, Faculdade de Ciências e Letras de Araraquara, Brazil, 2009.

CASALTA, M. *Análise de dados do CORDIAL-SIN O Sujeito de terceira pessoa do plural em PST e UNS*. Trabalho final de Dialetologia. Lisbon: FLUL, 2017. Análise de dados do CORDIAL-SIN O Sujeito de terceira pessoa do plural em PST e UNS (researchgate.net).

MARTINS, A. M. (coord.) [2000-]. CORDIAL-SIN: Corpus Dialectal para o Estudo da Sintaxe / Syntax-oriented Corpus of Portuguese Dialects. Lisboa, Centro de Linguística da Universidade de Lisboa. URL: http://www.clul.ulisboa.pt/en/10-research/314-cordial-s

CHOCIEJ, J. *Polish null subjects: English influence on heritage Polish in Toronto*, Unpublished ms, Department of Linguistics, University of Toronto, 2011.

CHOMSKY, N. *Lectures on Government and Binding*. Dordrecht: Foris, 1981.

CYRINO, S.; DUARTE, M.E.L.; KATO, M. A. Visible subjects and invisible clitics in Brazilian Portuguese. In: KATO, M. A.; NEGRÃO, E.V. (ed.), *Brazilian Portuguese and the Null Subject Parameter*. Vervuert: Frankfurt am Main, 2000, p.55-74.

D'ALESSANDRO, R. The Null Subject Parameter Where are we and where are we headed, 2014.

D'ALESSANDRO, R., Null Subjects. In: FÁBREGAS, A., MATEU, J. e PUTNAM, M. (ed.). *Contemporary Linguistic Parameters*. London: Bloomsbury Academic, 2015, p. 201-226.

DE CROUSAZ, I.; SHLONSKY, U. The Distribution of a Subject Clitic Pronoun in a Franco-Provençal Dialect and the Licensing of Pro, *Linguistic Inquiry,* XXXIV/3, 2003, p. 413-442.

DIÉMOZ, F. *Morphologie et syntaxe des pronoms personnels sujets dans les parlers francoprovençaux de la Vallée d'Aoste*. Tübingen: A. Francke, 2007.

DRYER, M. S. Expression of pronominal subjects. IN: DRYER, M.S.; HASPELMATH, M. (ed.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology, 2013.

DRYER, M.S.; HASPELMATH, M. (ed.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology, 2013 (Accessed on 18 Dec 2015).

DUARTE, M. E. L., A remarcação em curso no valor do parâmetro do sujeito nulo. *Cuadernos de la ALFAL*, 12 (2), p. 71-99, 2020.

DUARTE, M. E. L., Do pronome nulo ao pronome pleno: a trajetória do sujeito no português do Brasil In: ROBERTS, I., e KATO, M. A. (ed.) *Português brasileiro: Uma viagem diacrônica. Homenagem a Fernando Tarallo*. Editora da Unicamp, Campinas, Brazil, 1993, p. 107–129.

DUARTE, M. E.L. , Sociolinguistics as a powerful tool to follow the course of a parametric change, *University of Pennsylvania Working Papers in Linguistics*, Vol. 25 (2) , Article 3. 2020. Available at: https://repository.upenn.edu/pwpl/vol25/iss2/3

DUARTE, M.E.L. The loss of the avoid pronoun principle in Brazilian Portuguese. In: KATO, M.A.; NEGRÃO, E.V. (ed.), *Brazilian Portuguese and the null subject parameter*. Frankfurt am Main: Vervuert Verlag, 2000, p. 17–36.

DUARTE, M.E.L.; REIS, E. P. R. Revisitando o sujeito pronominal vinte anos depois. *ReVEL* 16(30), p.173–197, 2018.

DUARTE, M. E. L. *A Perda do Princípio 'Evite Pronome' no Português Brasileiro.* PhD thesis. Universidade de Campinas: São Paulo, Brazil, 1995.

FIGUEIREDO SILVA, M.C., O estado da arte dos estudos sobre sujeitos nulos, posições de sujeito e marcas flexionais. *Revista Linguística,* 13(2), p. 191-211, 2017. DOI: http://dx.doi.org/10.31513/linguistica.2017.v13n2a13510

FLORES, C. ; RINKE, E., Factors Constraining Subject Expression in European Portuguese Spoken in Hamburg. A Bi-Generational Corpus Investigation, *Discours*, 26 | 2020, http://journals.openedition.org/discours/10648.

GALVES, C.A sintaxe do português brasileiro. - *Ensaios de Lingüística* 7, 1987.

GILLIGAN, G.M. *A cross linguistic approach to the pro-drop parameter.* PhD dissertation: University of Southern California, 1987.

HASPELMATH, M. Parametric versus functional explanations of syntactic universals, In: BIBERAUER, T. (ed). *The Limits of Syntactic Variation*, Amsterdam: Benjamins, 2008, p. 75-107.

HEAP, D. La variation grammaticale en géolinguistique: Les pronoms sujet en roman central. München: Lincomm Europa, 2000.

HELASVUO, M-L. Agreement or crystallization: Patterns of 1st and 2nd person subjects and verbs of cognition in Finnish conversational interaction. *Journal of Pragmatics* 63, p. 63–78, 2014.

HENRIQUES, Y. *Particularidades morfossintáticas em variedades rurais do português falado na Ilha da Madeira*. Master thesis, University of Zurich, Switzerland, 2019.

HINZELIN, M-O. ; KAISER, G. A. Le paramètre du sujet nul dans les variétés dialectales de l'occitan et du francoprovençal, In: BARRA-JOVER, M (dir.), *Études de Linguistique gallo-romane*. Vincennes: Université de Vincennes, p. 247-260, 2012.

HOLMBERG, A. Null subject parameters. In: BIBERAUER, T., HOLMBERG, A., ROBERTS, I. e SHEEHAN, M., (ed.). *Parametric variation: null subjects in minimalist theory.* Cambridge: Cambridge University Press, 2010, p. 88–124.

HOLMBERG, A.; SHEEHAN, M. Control into finite clauses in partial null-subject languages, In: BIBERAUER, T., HOLMBERG, A., ROBERTS, I.; SHEEHAN, M., (ed.). *Parametric variation: null subjects in minimalist theory.* Cambridge: Cambridge University Press, 2010, p.125-152.

HOLMBERG, A; NAYUDU, A; SHEEHAN, M. Three partial null-subject languages: a comparison of Brazilian Portuguese, Finnish and Marathi. *Studia Linguistica*, 63(1), p. 59-97, 2009.

JAEGGLI, O.; SAFIR, K. (org.). *The Null Subject Parameter*. Kluwer Academic Publishers, 1989.

JIA, L.; BAYLEY, R. Null pronoun variation in Mandarin Chinese. *University of Pennsylvania Working Papers in Linguistics* 8(3), p.103–116, 2002

KAISER, G. A.; HACK, F. M. Sujets et sujets nuls en romanche, In: ILIESCU, M. *et al*. (ed.). *XXVe CILPR Congrès International de Linguistique et de Philologie Romanes, Tome 7, Section 15*, Berlin: De Gruyter Mouton, 2010, p. 83-91.

KATO, M. A.; DUARTE, M. E. L. Restrições na distribuição de sujeitos nulos no Português Brasileiro. *Veredas* 18(1), p. 1–24, 2014.

KATO, M.A. A evolução da noção de parâmetros. *D.E.L.T.A.,* 18:2, p. 309-337, 2002.

KATO, M.A. As controvérsias sobre o sujeito nulo no português brasileiro, In: CASTAGNA, V.; QUAREZEMIN, S. (org.). *Travessias em língua portuguesa Pesquisa linguística, ensino e tradução*. SAIL 16, Edizioni Ca'Foscari, 2020, p. 74-92. DOI 10.30687/978-88-6969-461-5/004.

KATO, M.A.; DUARTE, M.E.L. Restrições na distribuição de sujeitos nulos no Português Brasileiro, Veredas, 18(1), 2014. 24970-Texto do artigo-97931-1-10-20160616 (1).pdf

LABOV, W. Quantitative reasoning in linguistics. In: AMMON, U.; DITTMAR, N.; MATTHEIER, K. J.; TRUDGILL, P. (ed.). *Sociolinguistics/Soziolinguistik: An international handbook of the science of language and society/, Ein internationales Handbuch zur Wissenschaft von Sprache und Gesellschaft*, vol. 1, Berlin & New York: Mouton de Gruyter, 2005, p. 6–22.

LASTRA, Y.; BUTRAGUEÑO, P. M. Subject pronoun expression in oral Mexican Spanish. In: CARVALHO, A.M.; OROZCO, R.; SHIN, N. L. (ed.), *Subject pronoun expression in Spanish: A cross-dialectal perspective*, Washington, DC: Georgetown University Press, 2015, p. 39–57.

LEE, D-Y.; YONEZAWA, Y.The role of the overt expression of first and second person subject in Japanese. *Journal of Pragmatics* 40(4), p.733–767, 2008

LOBO, M. Sujeito nulo: Sintaxe e interpretação. In: RAPOSO, E.B.P.; NASCIMENTO, M.F.B. do; MOTA, M.A.C.da; SEGURA, L.; Mendes, A. (ed.) *Gramática do Português*. Lisbon: Fundação Calouste Gulbenkian, 2013, p. 2309-2335.

LOBO, M. Sujeitos nulos: gramática do adulto, aquisição de l1 e variação dialetal. In: MARTINS, A.M.; CARRILHO, E. (eds*). Manual de linguistica portuguesa*. Berlin: DE GRUYTER, 2016, p. 559-561. https://doi.org/10.1515/9783110368840-023

LOBO, M.; MARTINS, A. M. Subjects. In: DUFTER, A.; STARK, E. (eds). *Manual of Romance Morphosyntax and Syntax*. Berlin: De Gruyter, 2017, p. 27-88.

LUCCHESI, D. A realização do sujeito pronominal. In: LUCCHESI, D.; BAXTER, A.; RIBEIRO, I. (org.). *O português afro-brasileiro.* Salvador: EDUFBA, 2009, p. 165-183.

MADEIRA, A.; FIÉIS, A.; TEIXEIRA, J. Microvariação na resolução de sujeitos pronominais: português europeu vs. Italiano. *Revista da Associação Portuguesa de Linguística*, Nº8, p. 155-167, 2021. DOI: https://doi.org/10.26334/2183-9077/rapln8ano2021a11

MIESTAMO, M.; SINNEMÄKI, K.; KARLSSON, F. *Language Complexity: Typology, Contact, Change*. Amsterdam: John Benjamins Publishing Company, 2008.

MONTALBETTI, M. *After Binding Properties. On the Interpretation of Pronouns*. PhD thesis, Massachusetts, MA, USA: MIT, 1984.

MYAMOTO, A. M.; KAILER, D. Sujeito nulo no português brasileiro: variação estável ou mudança em curso? *A Cor das Letras* 20(1), p. 24-32, 2019. DOI: 10.13102/cl.v20i1.4622

NAGY, N. A sociolinguistic view of null subjects and VOT in Toronto heritage languages, *Lingua,* 164, p. 309-327, 2015.

NAGY, N.; AGHDASI, N.; DENIS, D.; MOTUT, A. Null subjects in heritage languages: Contact effects in a cross-linguistic contexto. *University of Pennsylvania Working Papers in Linguistics,* Vol.17 (2), Selected Papers from NWAV 39, 2011.

NURC. https://nurcrj.letras.ufrj.br.

OH, S-Y, Overt reference to speaker and recipient in Korean. *Discourse Studies* 9(4), p. 462–492, 2007. https://doi.org/10.1177%2F1461445607079163

OLIVEIRA, K. R. S. *Sujeito nulo ou sujeito preenchido? Uma análise da variedade falada pela comunidade quilombola Rio das rãs – Bahia*, dissertação de mestrado, Vitória da Conquista., Bahia, Brazil, 2020.

OLIVEIRA, M.; SANTOS, E. Pronomes nulos na posição de sujeito no português de Angola – um estudo preliminar. *Filol. lingüíst. port.*, 9, p. 85-101, 2007.

OLIVEIRA, V.M.S. *A expressão do sujeito no português de Moçambique*, Masters dissertation, UNICAMP, Brazil, 2016.

OLIVIÉRI, M. Paramètre du sujet nul et inversion du sujet dans les dialectes italiens et occitans. *Cahiers de Grammaire* 29, p. 105-120, 2004.

OLIVIÉRI, M.; LAI, J-P.; HEAP, D. Partial subject paradigms and feature geometry in Northern Occitan dialects. In: PERPIÑÁN, S.; HEAP, D.; MORENO-VILLAMAR, I.; SOTO-COROMINAS, A. (ed.). *Romance Languages and Linguistic Theory* 11: Selected papers from the 44th Linguistic Symposium on Romance Languages (LSRL), London, Ontario, 2017, p. 157-167. https://doi.org/10.1075/rllt.11.07oli

OROZCO, R. Pronominal variation in Colombian Costeño Spanish. In: CARVALHO, A.M.; OROZCO, R.; SHIN, N.L. (ed.). *Subject pronoun expression in Spanish: A cross-dialectal perspective*. Washington, DC: Georgetown University Press, 2015, p. 17–37.

OROZCO, R.; GUY, G. El uso variable de los pronombres sujetos: ¿Qué pasa en la costa Caribe colombiana?, In: WESTMORELAND, M.; THOMAS, J. *Selected Proceedings of the 4th Workshop on Spanish Sociolinguistics,* Somerville, MA: Cascadilla, 2008, p. 70-80.

OROZCO, R.; HURTADO, L.M. A variationist study of subject pronoun expression in Medellín, Colombia. *Languages 6: 5.* https://doi.org/ 10.3390/languages601005, 2021.

OTHEGUY, R.; ZENTELLA, A.C. *Spanish in New York: Language contact, dialect levelling, and structural continuity*. Oxford: Oxford University Press, 2012.

OTHEGUY, R.; ZENTELLA, A.C.; LIVERT, D. Language and dialect contact in Spanish of New York: Toward the formation of a speech community. *Language* 83(4), p. 770–802, 2007.

POLETTO, C., *La sintassi del soggetto nei dialetti italiani settentrionali.* Padova: Unipress, 1993.

POSIO, P. Properties of Pronominal Subjects. In:. GEESLIN, K. L. (ed)., *The Cambridge Handbook of Spanish Linguistics*. Cambridge: Cambridge University Press, 2018, p. 286-306.

POSIO, P. Subject expression in grammaticalizing constructions: The case of creo and acho 'I think' in Spanish and Portuguese. *Journal of Pragmatics* 63, p. 5–18, 2014.

POSIO, P. You and we: Impersonal second person singular and other referential devices in Spanish sociolinguistic interviews. *Journal of Pragmatics,* **99**, p. 1–16, 2016.

POSIO, P. Subject pronoun usage in formulaic sequences: Evidence from Peninsular Spanish In: CARVALHO, A.M.; OROZCO, R.; SHIN, N.L. (ed.). Subject pronoun expression in Spanish: A cross-dialectal perspective. Washington, DC: Georgetown University Press, 2015, p.59–78.

POSIO, P. The expression of first-person-singular subjects in spoken Peninsular Spanish and European Portuguese: Semantic roles and formulaic sequences. *Folia Linguistica* 47 (1), p. 253-292, 2013.

PRADA PÉREZ, A. de First person singular subject pronoun expression in Spanish in contact with Catalan. In: CARVALHO, A.M.; OROZCO, R.; SHIN, N.L. (ed.). *Subject pronoun expression in Spanish: A cross-dialectal perspective*.Washington, DC: Georgetown University Press, 2015, p. 121-142.

VERTENTES. http://www.vertentes.ufba.br/home.

REIS, E.P. R. *Atravessando o oceano: a expressão do sujeito pronominal de referência definida no português europeu insular.* Monography. Rio de Janeiro: UFRJ, Brasil, 2020.

RIZZI, L. On the status of subject clitics in Romance. In: JAEGGLI, O.; SILVA-CORVALÁN, C. (eds.). Studies in Romance linguistics. Dordrecht: Foris, 1986, p. 391-419.

RIZZI, L. *Issues in Italian syntax*. Dordrecht: Foris, 1982.

SAMPSON, G.; GIL, D.; TRUDGILL, P. (ed.). *Language Complexity as an Evolving Variable.* Oxford; New York: Oxford University Press, 2009.

SAVOIA, L. M.; MANZINI, R. Les clitiques sujets dans les variétés occitanes et francoprovençales italiennes. *Corpus*, 9, p. 165-189, 2010.

SAVOIA, L.M.; MANZINI, M. R. Lexicalization of 3rd person object clitics: clitic enclisis and clitic drop. In: D'ALESSANDRO, R.; LEDGEWAY, A.; ROBERTS, I. (ed.). *Syntactic Variation. The Dialects of Italy*. Cambridge: Cambridge University Press, 2010, p. 102–118.

SHEEHAN, M. Subjects, null subjects, and expletives. In: FISCHER, S.; GABRIEL, C. (ed.). *Manual of Grammatical Interfaces in Romance*. Berlin: De Gruyter, 2016, p. 329-362. https://doi.org/10.1515/9783110311860-014

SHIN, N. L. Grammatical complexification in Spanish in New York: 3sg pronoun expression and verbal ambiguity. *Language Variation and Change* 26(3), p. 303–330, 2014.

SIBILLE, J. *La passion de saint André,* édition *critique suivie d'une* étude *linguistique comparée*. Paris: Honoré Champion, 2007.

SIBILLE, J. Les clitiques sujets dans le parler occitan de Chiomonte et des Ramats (Italie). *Revue de Linguistique Romane* 76, 401–435, 2012.

SILVA-CORVALÁN, C. Otra mirada a la expresión del sujeto como variable sintáctica. In; MORENO, F. F., et. Al. (ed.), *Lengua, variación y contexto: Estudios dedicados a Humberto López Morales*, vol. 2. Madrid: Arco Libros, 2003, p. 849–860.

SOARES, E. C. *Anaphors in discourse: anaphoric subjects in Brazilian Portuguese*. PhD Thesis. University Sorbonne Paris City, Paris, France, 2017.

SOARES, E.C.; MILLER, P.; HEMFORTH, B. O Efeito de Características Semânticas e Discursivas no Uso de Sujeitos Nulos e Pronominais - Um Estudo Quantitativo dos Sujeitos de Terceira Pessoa no Português do Brasil. *D.E.L.T.A.,* 36 (1), p. 1-38, 2020. https://doi.org/10.1590/1678-460X2020360107

TEIXEIRA, E. P., A representação do sujeito pronominal no português popular angolano, *PAPIA*, 22 (1), p. 141-159, 2012.

TORIBIO, J. A., Setting parametric limits on dialectal variation in Spanish. *Lingua* 110, p. 315–341, 2000.

TORRES CACOULLOS, R.; TRAVIS, C. E. Variationist typology: Shared probabilistic constraints across (non-)null subject languages. *Linguistics* 57, p. 653–92, 2019.

TRAVIS, C.E. Genre effects on subject expression in Spanish: priming in narrative and conversation. *Language Variation and Change* 19 (2), p.101-135, 2007.

TRAVIS, C.E.; TORRES CACOULLOS, R. What do subject pronouns do in discourse? Cognitive, mechanical and constructional factors in variation. *Cognitive Linguistics* 23(4), p. 711–748, 2012.

VIEIRA, S. R.; BRANDÃO, S. F. CORPORAPORT: Variedades do Português em análise. Rio de Janeiro: Faculdade de Letras-UFRJ. www.corporaport.letras.ufrj.br.